



**Manchester  
Metropolitan  
University**

---

Edmonds, Bruce ORCID logoORCID: <https://orcid.org/0000-0002-3903-2507>, Le Page, Christophe, Bithell, Mike, Chattoe-Brown, Edmund, Grimm, Volker, Meyer, Ruth, Montañola-Sales, Cristina, Ormerod, Paul, Root, Hilton and Squazzoni, Flaminio (2019) Different Modelling Purposes. *Journal of Artificial Societies and Social Simulation (JASSS)*, 22 (3). p. 6. ISSN 1460-7425

---

**Downloaded from:** <https://e-space.mmu.ac.uk/623529/>

**Version:** Published Version

**Publisher:** University of Surrey

**DOI:** <https://doi.org/10.18564/jasss.3993>

**Usage rights:** Creative Commons: Attribution 4.0

Please cite the published version

<https://e-space.mmu.ac.uk>

---

# Different Modelling Purposes

**Bruce Edmonds<sup>1</sup>, Christophe Le Page<sup>2</sup>, Mike Bithell<sup>3</sup>, Edmund Chattoe-Brown<sup>4</sup>, Volker Grimm<sup>5</sup>, Ruth Meyer<sup>6</sup>, Cristina Montaña-Sales<sup>7</sup>, Paul Ormerod<sup>8</sup>, Hilton Root<sup>9</sup>, Flaminio Squazzoni<sup>10</sup>**



<sup>1</sup>Centre for Policy Modelling, Manchester Metropolitan University Business School, All Saints Campus Oxford Road, Manchester M15 6BH, United Kingdom

<sup>2</sup>CIRAD - UPR GREEN, Campus international de Baillarguet, TAC-47 / F, Montpellier Cedex 34398, France

<sup>3</sup>Department of Geography, University of Cambridge, Downing Place, Cambridge CB2 3EN, United Kingdom

<sup>4</sup>School of Media, Communication and Sociology, University of Leicester, Bankfield House, 132 New Walk, Leicester LE1 7JA, United Kingdom

<sup>5</sup>Department of Ecological Modelling, Helmholtz Centre for Environmental Research - UFZ, Permoserstr. 15, 04318 Leipzig, Germany

<sup>6</sup>Centre for Policy Modelling, Manchester Metropolitan University, Aytoun Street, Manchester M1 3GH, United Kingdom

<sup>7</sup>IQS School of Management, Universidad Ramon Llull, Via Augusta 390, Barcelona 08017, Spain

<sup>8</sup>Volterra Consulting, 135c Sheen Lane, London SW14 8AE, United Kingdom

<sup>9</sup>Schar School of Policy and Government, George Mason University, 3351 Fairfax Dr., MS 3B1 Arlington, Virginia 22201, United States

<sup>10</sup>Department of Social and Political Sciences, University of Milan, Via Conservatorio 7, 20122 Milan, Italy

Correspondence should be addressed to [bruce@edmonds.name](mailto:bruce@edmonds.name)

*Journal of Artificial Societies and Social Simulation* 22(3) 6, 2019

Doi: 10.18564/jasss.3993 Url: <http://jasss.soc.surrey.ac.uk/22/3/6.html>

Received: 06-09-2018 Accepted: 08-06-2019 Published: 30-06-2019

---

**Abstract:** How one builds, checks, validates and interprets a model depends on its ‘purpose’. This is true even if the same model code is used for different purposes. This means that a model built for one purpose but then used for another needs to be re-justified for the new purpose and this will probably mean it also has to be re-checked, re-validated and maybe even re-built in a different way. Here we review some of the different purposes for a simulation model of complex social phenomena, focusing on seven in particular: prediction, explanation, description, theoretical exploration, illustration, analogy, and social interaction. The paper looks at some of the implications in terms of the ways in which the intended purpose might fail. This analysis motivates some of the ways in which these ‘dangers’ might be avoided or mitigated. It also looks at the ways that a confusion of modelling purposes can fatally weaken modelling projects, whilst giving a false sense of their quality. These distinctions clarify some previous debates as to the best modelling strategy (e.g. KISS and KIDS). The paper ends with a plea for modellers to be clear concerning which purpose they are justifying their model against.

**Keywords:** Models, Purpose, Prediction, Explanation, Theory, Analogy

---

## Introduction

- 1.1 A common view of modelling is that one builds a ‘life-like’ reflection of some system, which then can be relied upon to act like that system. This is a correspondence view of modelling where the details in the model correspond in a roughly one-one manner<sup>1</sup> with those in the modelling target — as if the model were some kind of ‘picture’ of what it models. We suggest that this picture analogy is not helpful when it comes to the justification or judging of models<sup>2</sup>.
- 1.2 Rather, we suggest a more pragmatic approach, where models are judged as tools designed for specific purposes — the goodness of the model being assessed against how good it is for its declared purpose<sup>3</sup>. Models therefore have been characterized as “purposeful representation” (Starfield et al. 1990): we are not modelling systems per se, but only with respect to a particular purpose. This purpose can act as a kind of filter, or customer, which helps us decide: is this element of the real system possibly important for our purpose? Thus, a forest model designed for estimating the amount of timber production will, almost certainly, be very different

from a forest model focussing on biodiversity — and one would not expect them to be the same, despite the fact they are modelling the same thing.

- 1.3 Although a model designed for one purpose may turn out to be OK for another, it is more productive to use a tool designed for the job in hand. One may be able to use a kitchen knife for shaping wood, but it is much better to use a chisel. In particular, we argue that even when a model (or model component) turns out to be useful for more than one purpose it needs to be re-justified with respect to *each* of the claimed purposes *separately*. To extend the previous analogy: a tool with the blade of a chisel but the handle of a kitchen knife may satisfy some of the criteria for a tool to carve wood and some of the criteria for a tool to carve cooked meat, but not be much good for either purpose — it fails as either. If one *did* come up with a new tool that is good at both, this would be because it could be justified for each purpose separately.
- 1.4 In his paper ‘Why Model?’, Epstein (2008) lists 17 different reasons for making a model: from the abstract, ‘discover new questions’, to the practical ‘educate the general public’. This illustrates both the usefulness of modelling but also the potential for confusion. As Epstein points out, the power of modelling comes from making an informal set of ideas formal. That is, they are made precise using unambiguous code or mathematical symbols. This lack of ambiguity has huge benefits for the process of science, since it allows researchers to more rigorously explore the consequences of their assumptions and to share, critique, and improve models without transmission errors (Edmonds 2010). However, in many social simulation papers the purpose that the model was developed for, or more critically, the purpose under which it is being presented is often left implicit or confused. Maybe this is due to the prevalence of the ‘correspondence picture’ of modelling discussed above, maybe the authors optimistically conceive of their creations being useful in many different ways, or maybe they simply developed the model without a specific purpose in mind. However, regardless of the reason, the consequence is that readers do not know how to judge the model when presented. This can have the result that models might avoid proper judgement — demonstrating partial success in different ways with respect to a number of purposes, but not adequacy against any. The recent rise of standard protocols like ODD (Grimm et al. 2006, 2010; Müller et al. 2013) for describing published models may help improving this inconvenient situation: stating the purpose of the model constitutes the first part of the framework, but so far ODD does not provide specific categories of purposes, so that often the “Purpose” section of ODD model descriptions is not specific enough to enable a reader to decide how to judge it.
- 1.5 Our use of language helps cement this confusion: we talk about a ‘predictive model’ as if it is something in the code that makes it predictive (forgetting the role of the mapping between the model and what it predicts). Rather we are suggesting a shift from the code as a thing in itself, to code as a tool for a particular purpose. This marks a shift from programming, where the focus is on the nature and quality of the code, to modelling, where the focus is on the relationship of the behaviour of some code to what is being modelled. Using terms such as ‘explanatory model’ is OK, as long as we understand that this is shorthand for ‘a model which establishes an explanation’ etc.
- 1.6 Producing, checking and documenting code is labour intensive (Janssen 2017). As a result, we often wish to reuse some code produced for one purpose for another purpose. However, this often causes as much new work as it saves due to the effort required to justify code for a new purpose, and — if this extra work is not done — the risk that time and energy of many researchers are wasted due to the confusions that can result<sup>4</sup>. In practice, we have seen very little code that does not need to be re-written when one has a new purpose in mind<sup>5</sup>. Ideas can be transferred and well-honed libraries reused for very well defined purposes, but not the core code that makes up a model of complex social phenomena<sup>6</sup>. This is partly because different uses imply different assumptions and be open to different risks (as we will argue below).
- 1.7 Although one can have many possible purposes for a model, in this paper, we will look at seven. These are:
  - prediction,
  - explanation,
  - description,
  - theoretical exploration,
  - illustration,
  - analogy<sup>7</sup>,
  - and social learning<sup>8</sup>.

- 1.8** The first three of these are empirical — that is, they have a well-defined relationship with observed evidence<sup>9</sup>. The others are not directly related to what is observed, but concern ways of thinking about things, theoretical properties or communication. The point of this paper is to distinguish the different kinds of purpose, even if in common usage we may conflate them. The reason is that these different purposes imply very different ways of judging, justifying, checking, and even building models.
- 1.9** There are, of course, many other reasons for modelling (e.g. Epstein 2008) but the above seven seem to be the main ones that concern the papers in the field of social simulation (Squazzoni et al. 2014; Hauke et al. 2017). Given the pluralist and pragmatist intent of this paper, we have avoided trying to theorise too much as to how these purposes relate or cluster, but rather simply aim to distinguish them and discuss the more practical aspects that relate to each. There will be other purposes for modelling that we have not covered, or even thought of — this is fine, the main aim of the paper is to get authors to make their purpose crystal clear so that readers know on what basis their model should be judged in any presentation of their work.
- 1.10** Each of these purposes is discussed, in turn, below. For each purpose a brief ‘risk analysis’ is presented — some of the ways one might fail to achieve that purpose — along with some ways of mitigating these risks (which involve different activities, such as validation or verification<sup>10</sup>). In the penultimate section, some common confusions of purpose are illustrated and discussed, ending with a brief summary and plea to make one’s purpose clear. There is also a discussion of the role of modelling strategies and how these might relate to the different purposes but in order not to muddy our central point this is mostly relegated to an appendix. There are many endnotes, reflecting points the reviewers and authors want included, but you do not have to read them to get the sense of this paper.
- 1.11** In the discussion below, we focus on modelling within social simulation (broadly conceived) rather than modelling in general. Although we strongly suspect that very similar arguments could be made about modelling in other fields, we are not experts in those fields.
- 1.12** Our purpose for all this is to advance the culture, and practice, of social simulation modelling of all kinds. This kind of thing only works if it has direct benefits for the modeller adhering to it, which is the case here: by being explicit about the purpose of your model, its design will improve, as well as its reception and, possibly, further development.

## Prediction

### Motivation

- 2.1** If one can reliably predict anything that is not already known, this is undeniably useful regardless of the nature of the model (e.g. whether its processes are a reflection of what happens in the observed system or not<sup>11</sup>). For instance, the gas laws (stating e.g. that at a fixed pressure, the increase in volume of gas is proportional to the increase of temperature) were discovered long before the reason why they worked.
- 2.2** However, there is another reason that prediction is valued: it is considered the gold standard of science — the ability of a model or theory to predict is taken as the most reliable indicator of a model’s truth. This might be because this purpose leaves the least room for deceiving ourselves as to whether a model succeeds or fails — all other kinds are more amenable to tweaking the model until it is presentable.
- 2.3** Prediction is done in two principle ways: (a) model A fits the evidence better than model B — a comparative approach<sup>12</sup> or (b) model A is falsified (or not) by the evidence — a falsification approach. In either, the idea is that, given a sufficient supply of different models, better models will be gradually selected over time, either because the bad ones are discarded or outcompeted by better models.

### Definition

- 2.4** By ‘prediction’, we mean the ability to reliably anticipate well-defined aspects of data that is not currently known to a useful degree of accuracy via computations using the model.
- 2.5** Unpacking this definition:
- It has to do it *reliably* — that is under some known (but not necessarily precise) conditions the model will work; otherwise one would not know when one could use it.

- The data it anticipates has to be *unknown* to the modeller. ‘Predicting’ out-of-sample data is not enough, since pressures to re-do a model and get a better fit are huge and negative results are difficult to publish. It may be that a model tested on only known data does, in fact, predict well on unknown data, but this fact can not be demonstrated until it predicts unknown data. We should suspend judgement until it does.
- The aspects of the data that it predicts might be a numerical value, but also it might also be a pattern or a relationship (Grimm et al. 2005; Thorngate & Edmonds 2013; Watts 2014) — almost anything as long as this can be unambiguously checked to see if it holds. Talking about “aspects” is necessary due to the fact that data is almost never anticipated exactly, but only in some respects — e.g. taking into account noise, or only the distribution of the data.
- The anticipation has to be to a *useful* degree of accuracy. This will depend upon the purpose to which it is being put, e.g. as in weather forecasting. What counts as useful is a tricky subject and one that we do choose to discuss here, except to point out that it would be useful if authors specified what a useful degree of accuracy would be for their purposes.
- Sometimes a model can predict a completely unexpected pattern or outcome which can subsequently be empirically checked<sup>13</sup>. This is not the same as finding general phenomena that would seem to confirm the understanding suggested by a model — that is not prediction in our sense.

**2.6** Unfortunately, there are at least two different uses of the word ‘predict’ (Troitzsch 2009). Almost all scientific models ‘predict’ in the weak sense of being used to calculate or compute some result given some settings or data, but this is different from correctly anticipating *unknown* data<sup>14</sup>. For this reason, some use the term ‘forecast’ for anticipating unknown data and use the word ‘prediction’ for almost any inference of one aspect from another using a model<sup>15</sup>. However, this causes confusions in other ways so this does not necessarily make things clearer. *Firstly*, ‘forecasting’ implies that the unknown data is in the future (which is not always the case), *secondly*, large parts of science use the word ‘prediction’ for the process of anticipating unknown data, and *thirdly* it confuses those who use the words in an interchangeable manner. For example, if a modeller says their model ‘predicts’ something when they simply mean that it outputs it, then most of the audience may well assume the author is claiming more utility than is intended. Thus, we do not think this particular distinction is at all helpful. Others use prediction to include any kind of direct empirical relationship, whilst here we find it more helpful to distinguish between them.

**2.7** As Watts (2014) points out, useful prediction does not have to be a ‘point’ prediction of a future event. For example, one might predict that some particular thing will not happen, the existence of something in the past (e.g. the existence of Pluto), something about the shape or direction of trends or distributions (Thorngate & Edmonds 2013), comparative future performance (Klingert 2017) or even qualitative facts. The important fact is that what is being predicted is not known beforehand by the modeller, and that it can be unambiguously checked after it is known.

## An example

**2.8** Nate Silver is a forecaster, that is he aims to predict future social phenomena, such as the results of elections and the outcome of sports competitions. He gained fame when he correctly predicted the outcomes of all 50 electoral colleges in Obama’s election. This is a data-hungry activity, which involves the long-term development of simulations that carefully see what can be inferred from the available data. As well as making predictions, his unit tries to establish the level of uncertainty in those predictions — being honest about the probability of those predictions coming about given the likely levels of error and bias in the data. These models tend to be of a mostly statistical nature but can include elements of individual-based modelling (e.g. electoral colleges). As described in his book (Silver 2012) this involves a number of properties and activities, including:

- Repeated testing of the models against unknown data;
- Keeping the models fairly simple and transparent so one can understand clearly what they are doing (and what they do not cover);
- Encoding into the model aspects of the target phenomena that one is relatively certain about (such as the structure of the US presidential electoral college);
- Being heavily data-biased (as compared to theory-biased) in its method, that is requiring a lot of data to help eliminate sources of error and bias;

- Producing probabilistic predictions, giving a good idea about the level of uncertainty in any prediction;
- Being clear about what kind of factors are not covered in the model, so the predictions are relative to a clear set of declared assumptions and one knows the kind of circumstances in which one might be able to rely upon the predictions.

**2.9** Post hoc analysis of predictions — explaining why it worked or not — is kept distinct from the predictive models themselves — this analysis may inform changes to the predictive model but is not then incorporated into the model. The analysis is thus kept independent of the predictive model so it can be an effective check. Making a good predictive model requires a lot of time getting it wrong with real, unknown data, and trying again before one approaches qualified successful predictions.

## Risks

**2.10** Prediction (as we define it) is very hard for any complex social system. For this reason, it is rarely attempted<sup>16</sup>. Many re-evaluations of econometric models against data that has emerged since publication have revealed a high rate of failure. For example, Meese & Rogoff (1983) looked at 40 econometric models where they claimed they were predicting some time-series. However, 37 out of 40 models failed completely when tested on newly-available data from the same time series that they claimed to predict. Clearly, although presented as being predictive models, they did not actually predict *unknown* data. Many of these used the strategy of first dividing the data into in-sample and out-of-sample data, and then parameterising the model on the former and exhibiting the fit against the latter. Although we do not know for sure, presumably what happened was as follows. The apparent fit of the 37 models was not simply a matter of bad luck, but that all of these models had been (explicitly or implicitly) fitted to the out-of-sample data, because the out-of-sample data was known to the modeller before publication. That is, if the model failed to fit the out-of-sample data the first time the model was tested, it was then adjusted until it did work, or alternatively, only those models that fitted the out-of-sample data were published (a publishing bias). Thus, in these cases the models were not tested against predicting the out-of-sample data even though they were presented as such. Fitting known data is simply not a sufficient test for predictive ability<sup>17</sup>.

**2.11** There are many reasons why prediction of complex social systems fails<sup>18</sup>, but three of those specific to the social sciences are as follows: (1) it is unknown what processes are needed to be included in the model, (2) a lack of enough quality data of the right kinds and (3) having the right kind of data. We will discuss each of these in turn.

1. In the physical sciences, there are often well-validated micro-level models (e.g. fluid dynamics in the case of weather forecasting) that tell us what processes are potentially relevant at a coarser level and which are not. In the social sciences this is not the case — we do not know what the essential processes are. Here, it is often the case that there are other processes that the authors have not considered that, if included, would completely change the results. This is due to two different causes: (a) we simply do not know much about how and why people behave in different circumstances and (b) different limitations of intended context will mean that different processes are relevant.
2. Unlike in the physical sciences, there has been a paucity of the kind of data we would need to check the predictive power of models. This paucity can be due to (a) there is not enough data (or data from enough independent instances) to enable the iterative checking and adapting of the models on new sets of unknown data each time we need to, or (b) the data is not of the right kind to do this. What can often happen is that one has partial sets of data that require some strong assumptions in order to compare against the predictions in question (e.g. the data might only be a proxy of what is being predicted, or you need assumptions in order to link sets of data). In the former case, (a), one simply has not enough to check the predictive power in multiple cases so one has to suspend judgement as to whether the model predicts in general, until the data is available. In the latter case, (b), the success at prediction is relative to the assumptions made to check the prediction.
3. There is a further point. We may appear to have sufficient data of what seems to be the right kind. However, the data may contain relatively small amounts of true information. For example, Ormerod & Mounfield (2000) show, using modern signal processing methods, that data series in macroeconomics are dominated by noise rather than by information. They suggest that this is a key explanation for the very poor forecasting record in macroeconomics.

**2.12** A more subtle risk is that the conditions under which one can rely upon a model to predict well might not be clear. If this is the case then it is hard to rely upon the model for prediction in a new situation, since one does not know its conditions of application — i.e. *when* it predicts well<sup>19</sup>.



## Mitigating measures

**2.13** To ensure that a model does indeed predict well, one can seek to ensure the following:

- That the model has been tested on several cases where it has successfully predicted data unknown to the modeller (at the time of prediction);
- That information about the following are included: exactly what aspects it predicts, guidelines on when the model can be used to predict and when not, some guidelines as to the degree or kind of accuracy it predicts with, any other caveats a user of the model should be aware of;
- That a clear distinction between tweaking (calibration, model selection) and independent prediction is made.
- That the model code is distributed so others can explore when and how well it predicts.

## Explanation

### Motivation

- 3.1** Often, especially with complex social phenomena, one is particularly interested in understanding *why* something occurs — in other words, explaining it. Even if you cannot predict something before it is known, you still might be able to explain it afterwards. This distinction mirrors that in the physical sciences where there are both *phenomenological* as well as *explanatory laws* (Cartwright 1983) — the former match the data, the latter explain why they came about. Often we have good predictive or good explanatory models but not both. For example, the gas laws that link measurements of temperature, pressure and volume were known before the explanation in terms of molecules of gas bouncing randomly around, and similarly Darwin's explanatory Theory of Natural Selection was known long before the Neo-Darwinist synthesis with genetics that allowed for predictive models<sup>20</sup>.
- 3.2** Understanding is important for managing complex systems as well as for understanding when predictive models might work. Whilst generally with complex social phenomena explanation is easier than prediction, sometimes prediction comes first (however if one can predict then this invites research to explain *why* the prediction works).
- 3.3** If one makes a simulation in which certain mechanisms or processes are built in, and the outcomes of the simulation match some (known) data, then this simulation can support an explanation of the data using the built-in mechanisms. The explanation itself is usually of a more general nature and the traces of the simulation runs are examples of that account. Simulations that involve complicated processes can thus support complex explanations — that are beyond natural language reasoning to follow. The simulations make the explanation explicit, even if we cannot fully comprehend its detail. The formal nature of the simulation makes it possible to test the conditions and cases under which the explanation works, and to better its assumptions.

### Definition

- 3.4** By 'explanation' we mean establishing a possible causal chain from a set-up to its consequences in terms of the mechanisms in a simulation.
- 3.5** Unpacking some parts of this:
- The *possible causal chain* is a set of inferences or computations made as part of running the simulation — in simulations with random elements, each run will be slightly different. In this case, it is either a possibilistic explanation (A could cause B), so one just has to show one run exhibiting the complete chain, or a probabilistic explanation (A probably causes B, or A causes a distribution of outcomes around B), in which case one has to look at an assembly of runs, maybe summarising them using statistics or visual representations.
  - For explanatory purposes, the structure of the model is important, because that limits what the explanation consists of. For example, if social norms were a built-in mechanism in a simulation and it resulted in some cooperation, then that cooperation could be explained (at least partially) in terms of social norms.

If, on the other hand, the model consisted of mechanisms that are known *not* to occur, any explanation one established would be in terms of these non-existent mechanisms — which is not very helpful. If one has parameterised the simulation on some in-sample data (found the values of the free parameters that made the simulation fit the in-sample data) then the explanation of the outcomes is also in terms of the in-sample data, mediated by these free parameters<sup>21</sup>.

- The *consequences* of the simulations are generally measurements of the outcomes of the simulation. These are compared with the data to see if it ‘fits’. It is usual that only some of the aspects of the target data and the data the simulation produces are considered significant — other aspects might not be (e.g. might be artefacts of the randomness in the simulation or other factors extraneous to the explanation). The kind of fit between data and simulation outcomes needs to be assessed in a way that is appropriate to which aspects of the data are significant and which are not. For example, if it is the level of the outcome that is key then a distance or error measure between this and the target data might be appropriate, but if it is the shape or trend of the outcomes over time that is significant then other techniques will be more appropriate (e.g. Thorngate & Edmonds 2013).

## Example

- 3.6** Stephen Lansing spent time in Bali as an anthropologist, researching how the Balinese coordinated their water usage (among other things). He and his collaborator, James Kramer, build a simulation to show how the Balinese system of temples acted to regulate water usage, through an elaborate system of agreements between farmers, enforced through the cultural and religious practices at those temples (Lansing & Kremer 1993). Although their observations could cover many instances of localities using the same system of negotiation over water, they were necessarily limited to all their observations being within the same culture. Their simulation helped establish the nature and robustness of their explanation by exploring a close universe of ‘what if’ questions, which vividly showed the comparative advantages of the observed system that had developed over a considerable period. The model does not predict that such systems will develop in the same circumstances, but it substantially adds to the understanding of the observed case.

## Risks

- 3.7** Clearly, there are several risks in the project of establishing a complex explanation using a simulation — what counts as a good explanation is not as clear-cut as what is a good prediction.
- 3.8** *Firstly*, the fit to the target data to be explained might be a very special case. For example, if many other parameters need to have very special values for the fit to occur, then the explanation is, at best, brittle and, at worst, an accident.
- 3.9** *Secondly*, the process that is unfolded in the simulation might be poorly understood so that the outcomes might depend upon some hidden assumption encapsulated in the code. In this case, the explanation is dependent upon this assumption holding, which is problematic if this assumption is very strong or unlikely.
- 3.10** *Thirdly*, there may be more than one explanation that fits the target data. So, although the simulation establishes one explanation it does not guarantee that it is the only candidate for this.

## Mitigating measures

- 3.11** To improve the quality and reliability of the explanation being established:
- Ensure that the mechanisms built into the simulation are plausible, or at least relate to what is known about the target phenomena in a clear manner;
  - Be clear about which aspects of the outcomes are considered significant in terms of comparison to the target data — i.e. exactly which aspects of that target data are being explained;
  - To reduce the risk of getting the right results for the wrong reasons, try to get the model reproduce multiple patterns simultaneously (e.g. observed at different levels of organisations and different scales). Each pattern serves as a filter of unrealistic parameter values and process representations. While single filters can be inefficient, for example cyclic dynamics, which are easy to generate, entire sets of patterns, usually 3-5, are much more difficult to reproduce. Thereby the risks of tweaking are not eliminated, but reduced (Grimm et al. 2005, 2010; Railsback & Grimm 2011);



- Probe the simulation to find out the conditions for the explanation holding using sensitivity analysis, addition of noise, multiple runs, changing processes not essential to the explanation<sup>22</sup> to see if the results still hold, and documenting assumptions;
- Do experiments in the classic way, to check that the explanation does, in fact, hold for your simulation code — i.e. check your code and try to refute the explanation using carefully designed experiments with the model.

## Description

### Motivation

- 4.1** An important, but currently under-appreciated, activity in science is that of description. Charles Darwin spent a long time sketching and describing the animals he observed on his travels aboard the HMS Beagle. These descriptions and sketches were not measurements or recordings in any direct sense, since he was already selecting from what he perceived and only recording an abstraction of what he thought of as relevant. Later on, these were used to illustrate and establish his theoretical abstraction — his theory of evolution of species by natural selection. The purpose of the descriptions were to record, in a coherent way, a set of selected aspects of the phenomena under observation. These might be used to inform later models with a different purpose or, if there are many of them, be the basis from which induction is done.
- 4.2** One can describe things using natural language, or pictures, but these are inadequate for dynamic and complex phenomena, where the essence of what is being described is how several mechanisms might relate over time. An agent-based simulation framework allows for a direct representation (one agent for one actor) without theoretical restrictions. It allows for dynamic situations as well as complex sets of entities and interactions to be represented (as needed). This can make it an ideal complement to scenario development because it ensures consistency between all the elements and the outcomes. It is also a good base for future generalisations when the author can access a set of such descriptive simulations.

### Definition

- 4.3** *A description (using a simulation) is an attempt to partially represent what is important of a specific observed case (or small set of closely related cases).*
- 4.4** Unpacking some of this:
- This is not an attempt to produce a 1-1 representation of what is being observed but only of the features thought to be relevant for the intended kind of study. It will leave out some features, however a description tends to err on the side of including aspects rather than excluding them (for simplicity, communication etc.).
  - It is not in any sense general, but seeks to capture a restricted set of cases — it is specific to these and no kind of generality beyond these can be assumed.
  - The simulation has to relate in an explicit and well-documented way to a set of evidence, experiences and data. This is the opposite of theoretical exposition and should have a direct and immediate connection with observation, data or experience.
  - This kind of purpose may involve the integration, within a single picture, of many different kinds of evidence, including: qualitative, expert opinion, social network data, survey data, geographic data, and longitudinal data.

### Examples

- 4.5** In 1993, François Bousquet and colleagues (Bousquet et al. 1993) developed a simulation<sup>23</sup> whose purpose was described as follows:

*“We have developed a simulator to represent both social, economic and ecological knowledge to contribute to a synthesis of the multidisciplinary knowledge — As a result of the simulations, focus can be put on the relation between space sharing rules and the evolution of the ecological equilibrium. The simulator is considered as a discussion tool to lead to interdisciplinary meetings.”*

- 4.6** Thus this is not designed to explain or predict but to represent and formalise, aiming to formalise some of the relationships and dynamics of the situation.
- 4.7** Moss (1998) describes a model that captures some of the interactions in a water pumping station during crises. This came about through extensive discussions with stakeholders within a UK water company about what happens in particular situations during such crises. The model sought to directly reflect this evidence within the dynamic form of a simulation, including cognitive agents who interact to resolve the crisis. This simulation captured aspects of the physical situation, but also tackled some of the cognitive and communicative aspects. To do this, he had to represent the problem solving and learning of key actors, so he inevitably had to use some existing theories and structures — namely, Alan Newell and Herbert Simon’s general problem solving architecture (Newell & Simon 1972) and Cohen’s endorsement mechanism (Cohen 1984). However, this is all made admirably explicit in the paper. The paper is suitably cautious in terms of any conclusions, saying that the simulation

*“indicate[s] a clear need for an investigation of appropriate organizational structures and procedures to deal with full-blown crises.”*

## Risks

- 4.8** Any system for representation will have its own affordances — it will be able to capture some kinds of aspect much more easily than others will. This inevitably biases the representations produced, as those elements that are easy to represent are more likely to be captured than those which are more difficult. Thus, the medium will influence what is captured and what is not.
- 4.9** Since agent-based simulation is not theoretically constrained<sup>24</sup>, there are a large number of ways in which any observed phenomena could be expressed in terms of simulation code. Thus, it is almost inevitable that any modeller will use some structures or mechanisms that they are familiar with in order to write the code<sup>25</sup>. Such a simulation is, in effect, an abduction with respect to these underlying structures and mechanisms — the phenomena are seen through these and expressed using them.
- 4.10** Finally, a reader of the simulation may not understand the limitations of the simulation and make false assumptions as to its generality. In particular, the inference within the simulations may not include all the processes that exist in what is observed — thus it cannot be relied upon to either predict outcomes or justify any specific explanation of those outcomes.

## Mitigating measures

- 4.11** As long as the limitations of the description (in terms of its selectivity, inference and biases) are made clear, there are relatively few risks here, since not much is being claimed. If it is going to be useful in the future as part of a (slightly abstracted) evidence base, then its limitations and biases do need to be explicit. The data, evidence or experience it is based upon also needs to be made clear. Thus, good documentation is the key here — one does not know how any particular description will be used in the future so the thoroughness of this is key to its future utility. Here it does not matter if the evidence is used to specify the simulation or to check it afterwards in terms of the outcomes, all that matters is that the way it relates to evidence is well documented. Standards for documentations (such as the ODD and its various extensions (Grimm et al. 2006, 2010) help ensure that all aspects are covered.

## Theoretical Exposition

### Motivation

- 5.1** If one has a mathematical model, one can do analysis upon its mathematics to understand its general properties. This kind of analysis is both easier and harder with a simulation model; to find out the properties of simulation code one just has to run the code — but this just gives one possible outcomes from one set of initial parameters. Thus, there is the problem that the runs one sees might not be representative of the behaviour in general. With complex systems, it is not easy to understand how the outcomes arise, even when one knows the full and correct specification of their processes, so simply knowing the code is not enough. Thus with highly

complicated processes, where the human mind cannot keep track of the parts unaided, one has the problem of understanding how these processes unfold in general.

- 5.2** Where mathematical analysis is not possible, one has to explore the theoretical properties using simulation — this is the goal of this kind of model. Of course, with many kinds of simulation one wants to understand how its mechanisms work, but here this is the only goal: to rigorously explore the consequences of assumptions, using mathematics and computer simulation. Thus, this purpose could be seen as more limited than the others, since some level of understanding the mechanisms is necessary for the other purposes (except maybe black-box predictive models). However, with this focus on just the mechanisms, there is an expectation that a more thorough exploration will be performed — *how* these mechanisms interact and when they produce different kinds of outcome.
- 5.3** Thus, the purpose here is to give some more general idea of how a set of mechanisms work, so that modellers can understand them better when used in models for other purposes. If the mechanisms and exploration were limited it would greatly reduce the usefulness of doing this. General insights are what is wanted here.
- 5.4** In practice, this means a mixture of inspection of data coming from the simulation, experiments and maybe some inference upon or checking of the mechanisms. In scientific terms, one makes a hypothesis about the working of the simulation — why some kinds of outcome occur in a given range of conditions — and then tests that hypothesis using well-directed simulation experiments.
- 5.5** The complete set of simulation outcomes over all possible initialisations (including random seeds) does encode the complete behaviour of simulation, but that is too vast and detailed to be comprehensible. Thus, some general truths covering the important aspects of the outcomes under a given range of conditions is necessary — the complete and certain generality established by mathematical analysis might be infeasible with many complex systems but we would like something that approximates this using simulation experiments.

## Definition

- 5.6** *‘Theoretical exposition’ means establishing then characterising (or assessing) hypotheses about the general behaviour of a set of mechanisms (using a simulation).*
- 5.7** Unpacking some key aspects here.
- One may well spend some time illustrating the discovered hypothesis (especially if it is novel or surprising), followed by a sensitivity analysis, but the crucial part is showing the hypotheses are refuted or not by a sequence of simulation experiments.
  - The hypotheses need to be (at least somewhat) general to be useful.
  - A use for theoretical exposition can be to refute a hypothesis, by exhibiting a concrete counter-example, or to establish a hypothesis. Whilst some theoretical exposition work involves mapping the outcomes systematically, it is nice to have explicit hypotheses about the results formulated and then tested with specific experiments designed to try and falsify them.
  - Another use of theoretical exposition is when assumptions are tested or compared — for example by weakening or changing an assumption or mechanism in an existing model and seeing if the outcomes change in significant ways from the original<sup>26</sup>.
  - Although any simulation has to have some meaning for it to be a model (otherwise it would just be some arbitrary code) this does not involve any other relationship with the observed world in terms of a defined mapping to data or evidence.

## Example

- 5.8** Deffuant et al. (2002) study the behaviour of a class of continuous opinion dynamic models. This paper does multiple runs, mapping the outcomes over the parameter space using some key indicators (e.g. relative mean influence of ‘extremist agents’). From these it summarises the overall behaviour of this class in qualitative terms, dividing the outcomes into a number of categories. Later Deffuant & Weisbuch (2007) use probability distribution models to try and study the behaviour of a class of continuous opinion dynamic models, approximating the agent-based models using equations. These analytic models were then tested against the agent-based models.

Although the work is vaguely motivated with reference to observed phenomena, the described research is to understand the overall behaviour of a class of models (Flache et al. 2017).

## Risks

- 5.9** In theoretical exposition one is not relating simulations to the observed world, so it is fundamentally an easier and ‘safer’ activity<sup>27</sup>. Since a near complete understanding of the simulation behaviour is desired this activity is usually concerned with relatively simple models. However, there are still risks — it is still easy to fool oneself with one’s own model. Thus, the main risk is that there is a bug in the code, so that what one thinks one is establishing about a set of mechanisms is really about a different set of mechanisms (i.e. those including the bug).
- 5.10** A second area of risk lies in a potential lack of generality, or ‘brittleness’ of what is established. If the hypothesis is true, but only holds under very special circumstances, then this reduces the usefulness of the hypothesis in terms of understanding the simulation behaviour.
- 5.11** Lastly, there is the risk of over-interpreting the results in terms of saying anything about the observed world. The model might suggest a hypothesis about the observed world, but it does not provide any level of empirical support for this.

## Mitigating measures

- 5.12** The measures that should be taken for this purpose are quite general and maybe best understood by the community of simulators.
- One needs to check one’s code thoroughly (see Galán et al. 2017 for a review of techniques).
  - One needs to be precise about the code and its documentation — the code should be made publicly available. This might include comments in the code, ODD style documentation in addition to a narrative account, and how the code relates to the theoretical assumptions.
  - Be clear as to the nature and scope of the hypotheses established.
  - A very thorough sensitivity check should be done, trying various versions with extra noise added, testing for extreme conditions etc.
  - It is good practice to illustrate the simulation so that the readers understand its key behaviours but then follow this with a series of attempted refutations of the hypotheses about its behaviour to show its robustness.
  - Be very careful about not claiming that this says anything about the observed world.

## Illustration

### Motivation

- 6.1** Sometimes one wants to make an idea clear or show it is possible, then an illustration is a good way of doing this. It does this by exhibiting a concrete example that might be more readily comprehended. Complex systems, especially complex social phenomena, can be difficult to describe, including multiple independent and interacting mechanisms and entities. A well-crafted simulation can help people see these complex interactions at work and hence appreciate these complexities better. As with description, this purpose does not support any significant claims. If the theory is already instantiated as a simulation (e.g. for theoretical exposition or explanation) then the illustrative simulation might well be a simplified version of this.
- 6.2** Illustration goes to the heart of the power of formal modelling, since it is the instantiation of a set of ideas in a structure that can be indefinitely inspected and critiqued. This example can then be used to make precise the meaning of the terms used to express that idea. It also shows the possibility of the process being illustrated (but nothing about its likelihood). One powerful use of an illustration is as a counter-example to a set of assumptions.

## Definition

**6.3** *An illustration (using a simulation) is to communicate or make clear an idea, theory or explanation.*

**6.4** Unpacking this.

- Here the simulation does not have to fully express what it is illustrating, it is sufficient that it gives a simplified example. So it may not do more than partially capture the idea, theory or explanation that it illustrates, and it cannot be relied upon for the inference of outcomes from any initial conditions or set-up.
- The clarity of the illustration is of over-riding importance here, not its veracity or completeness.
- An illustration should not make any claims, even of being a description. If it is going to be claimed that it is useful as a theoretical exposition, explanation or other purpose then it should be justified using those criteria — that it seems clear to the modeller is not enough.

## Example

**6.5** Schelling developed his famous model for a particular purpose — he was advising the Chicago district on what might be done about the high levels of segregation there. The assumption was that the sharp segregation observed must be a result of strong racial discrimination by its inhabitants. Schelling's model (Schelling 1969, 1971) showed that segregation could result from just weak preferences of inhabitants for their own kind — that even a wish for 30% of people of the same trait living in the neighbourhood could result in segregation. This was not obvious without building a model. The model was a clear counter example to the assumption.

**6.6** What the model did not do is say anything about what actually caused the segregation in Chicago — it might well be the result of strong racial prejudice — the model was not empirically connected with the case to say anything about this. What it did was illustrate an ideal about how segregation might occur. The model did not predict anything about the level of segregation, nor did it explain it. All it did was provide a counter-example to the current theories as to the cause of the segregation, showing that this was not *necessarily* the case<sup>28</sup>.

## Risks

**6.7** The main risk here is that you might deceive people using the illustration into reading more into the simulation than is intended. Just because a model's outputs mimics an observed pattern (e.g. a Lotka-Volterra cycle) there is no indication how much this model is relevant to similar examples of those patterns. One illustration is simply not enough to show a useful empirical relationship. There is also a risk of confusion if it is not clear which aspects are important to the illustration and which are not. A simulation for illustration will show the intended behaviour, but (unlike when its theory is being explored) it has been tested only for a restricted range of possibilities, indeed the claimed results might be quite brittle to insignificant changes in assumption.

## Mitigating measures

**6.8** Be very clear in the documentation that the purpose of the simulation is for illustration only, maybe giving pointers to fuller simulations that might be useful for other purposes. Also be clear in precisely what idea is being communicated, and so which aspects of the simulation are relevant for this purpose.

## Analogy

### Motivation

**7.1** Analogy is a powerful way of thinking about things. Roughly, it applies ideas or a structure from one domain and projects it onto another (Hofstadter & FARG (The Fluid Analogies Research Group) 1995). This can be especially useful for thinking about new or unfamiliar phenomena or as a guide to the direction of more rigorous thought. Sometimes it can result in new insights that are later established in a more rigorous way. There is some evidence that analogical thinking is very deeply entrenched in the whole way we think and communicate (Lakoff 2008). Almost anything can be used as an analogy, including simulation models.

**7.2** Playing about with simulations in a creative but informal manner can be very useful in terms of informing the intuitions of a researcher. In a sense, the simulation has illustrated an idea to its creator<sup>29</sup>. One might then exhibit a version of this simulation to help communicate this idea to others. How humans do analogy is not very well understood, but this does not mean that the simulation achieves any of the other purposes described above, and it is thus doubtful whether that idea has been established to be of public value (justifying its communication in a publication) until this happens. This is not to suggest that analogical thinking is not an important process in science. Providing new ways of thinking about complex mechanisms or giving us new examples to consider is a very valuable activity. However, this does not imply its adequacy for any other purpose.

### Definition

**7.3** *An analogy (using a simulation) is when processes illustrated by a simulation are used as a way of thinking about something in an informal manner.*

**7.4** Unpacking this.

**7.5** Figures 1, 2 and 3 illustrate the difference between common sense, scientific (i.e. empirical) and analogical understanding, respectively. The difference between empirical and analogical relationships is that in the latter there is no direct relationship between the analogy (in this case the simulation) and the evidence, but only indirectly through one's intuitive understanding. An empirical relationship may be staged (model  $\Leftrightarrow$  data  $\Leftrightarrow$  observations) but these are well-defined and testable.

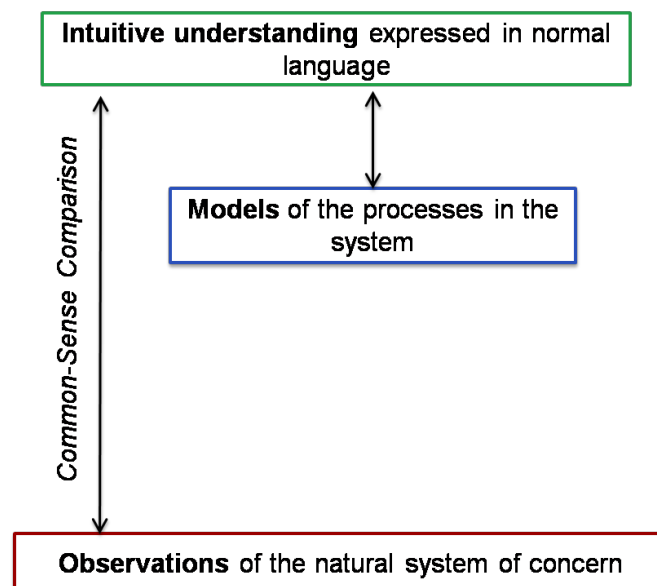


Figure 1: Common-sense mappings.



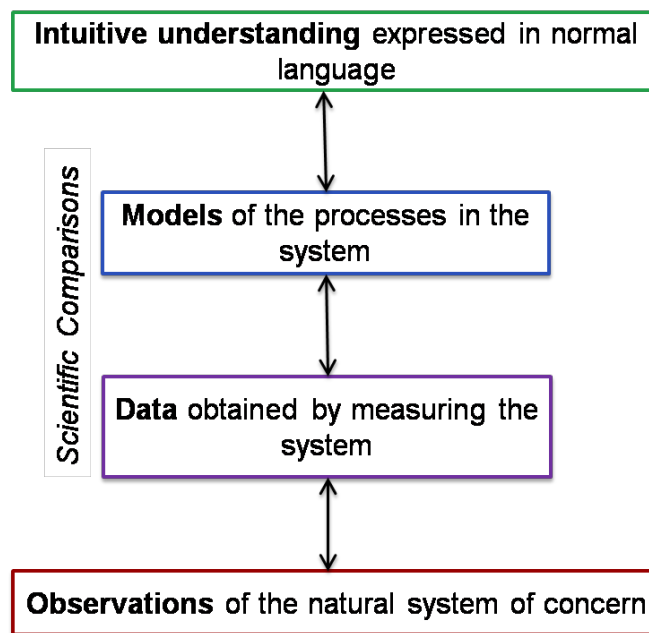


Figure 2: Scientific mappings.

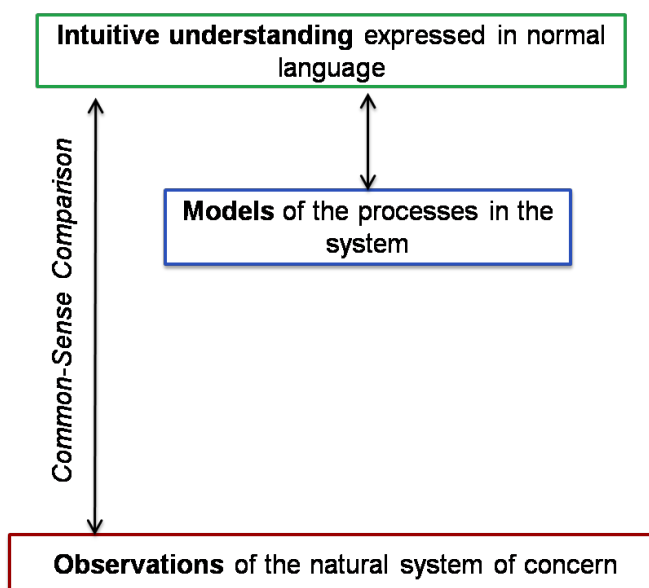


Figure 3: Analogical mappings.

- 7.6** It is not fully understood what the mind does when using an analogy to think, but it does involve taking the structure of some understanding from one domain and applying it to another. Humans have an innate ability to think using analogies and often do it unconsciously. Here the relationship between simulation and the observed is re-imagined for each domain in which it is applied<sup>30</sup> — thus this relationship is not well defined but flexible and creative.
- 7.7** The modelling purposes of analogy and illustration can be confused due to the fact that a simulation that illustrates an idea can later be used as an analogy (applying that idea to a new domain). However, the purposes *are* different — illustration is to make an idea clear and give a specific example of it; an analogy is when an idea is applied to other domains in an informal manner. Thus, there are substantive differences. For example, an illustrated idea might later be then tested in an explanatory model in a well-defined manner against some data, since how the illustration represents is precise. To take another example, an analogy might go well beyond that required of an illustration in terms of mapping many possibilities but without any well-defined relationship

with the real world.

#### 7.8 Specifically:

- An analogy might not be clear or well defined but an illustration should be;
- An illustration can be a single instance or example but an analogy might cover a whole class of processes or phenomena (albeit informally);
- A good analogy suggests a useful way of thinking about things, maybe producing new insights, a good illustration does not have to do this;
- An illustration can be used as a counter example to some widely held assumptions, whilst an analogy cannot.

### Example

- 7.9 In his book, Axelrod (1984) describes a formalised computational ‘game’ where different strategies are pitted against each other, playing the iterated prisoner’s dilemma. Some different scenarios are described, where it is shown how the tit for tat strategy can survive against many other mixes of strategies (static or evolving). The conclusions are supported by some simple mathematical considerations, but the model and its consequences were not explored in any widespread manner<sup>31</sup>. In the book, the purpose of the model is to open up a new way of thinking about the evolution of cooperation. The book claims the idea ‘explains’ many observed phenomena<sup>32</sup>, but in an analogical manner — no precise relationship with any observed measurements is described. There is no validation of the model here or in the more academic paper that described these results (Axelrod & Hamilton 1981). In the academic paper there are some mathematical arguments which show the plausibility of the model but the paper, like the book, progresses by showing the idea is coherent with some reported phenomena — but it is the ideas rather than the model that are so related. Thus, in this case, the simulation model is an analogy to support the idea, which is related to evidence in a qualitative manner — the relationship of the model to evidence is indirect (Edmonds 2001). Thus, the role of the simulation model is that of a way of thinking about the processes concerned and the model does not qualify for either explaining specific data, predicting anything unknown or exploring a theory.

### Risks

- 7.10 These simulations are just a way of thinking about other phenomena. However, just because you can think about some phenomena in a particular way does not make it true. The human mind is good at creating, ‘on the fly’, connections between an analogy and what it is considering — so good that it does it almost without us being aware of this process. The danger here is of confusing being able to think of some phenomena using an idea, and that idea having any force in terms of a possible explanation or method of prediction. The apparent generality of an analogy tends to dissipate when one tries to specify precisely the relationship of a model to observations, since an analogy has a different set of relationships for each situation it is applied to — it is a supremely flexible way of thinking. This flexibility means that it does not work well to support an explanation or predict well, since both of these necessitate an explicit and fixed relationship with observed data.

### Mitigating measures

- 7.11 The obvious mitigation is to make clear when one is using a model as an analogy and that no rigorous relationship with evidence is intended. Analogies can be a very productive way of thinking as long as one does not infer anything more from them — analogies can be very misleading and are not firm foundations from which inference can be safely made.

# Social Learning

## Motivation

- 8.1 Sometimes our simulations are not about the observed world at all, but (at least substantially) designed to reflect an actor's view of that world. Such simulations can be developed collectively by a group of people. In this case, the model can act as a mediator between members of the group and can result in a shared understanding of the world (or at least a clear idea where the members agree and differ). This shared understanding is explicitly encapsulated in the model for all to discuss, avoiding the misunderstandings that can come from more abstract discussion.
- 8.2 When the participatory dimension of the modelling process is to the fore, the overriding purpose can be referred to as promoting social learning, e.g. increasing the knowledge of individuals through social learning and processes within a social network and favouring the acquisition of collective skills (Reed et al. 2010). To build and use models in a collaborative way is here a means for: generating a shared level of information among participants, to create common knowledge, explore common goals, and understand the views, interests, and rationale of opposing parties about the target system. Approaches, such as mediated modelling (Van den Belt 2004) and companion modelling (Barreteau et al. 2003; Etienne 2014), have been applied to pursuing this goal in the domain of environmental sciences, and more specifically to tackle problems related to the collective management of renewable resources. As pointed out by Elinor Ostrom (2009), when users of a given socio-ecological system (SES) share common knowledge of relevant SES attributes as well as how their actions affect each other, they can more easily self-organise. There is here a radical shift from the purpose of using models for prediction for the purpose of seeking an optimal solution within a command-and-control structure, to the purpose of collective learning (Brugnach 2010), related to the concept of adaptive co-management which leitmotiv is *learning to manage and managing to learn* (Olsson et al. 2004).

## Definition

- 8.3 *A simulation is a tool for social learning when it encapsulates a shared understanding (or set of understandings) of a group of people.*
- 8.4 Unpacking this.
- 8.5 Developing a model as a tool for social learning is distinct from when one uses people as a source of evidence in the form of their expert or personal opinion for a descriptive or explanatory model. In this case, although the source is human and might be subjectively biased, the purpose is to produce an objectively justifiable model. Although a model produced as part of a participatory social learning process might have other uses, these are not what such a process is designed for.
- 8.6 Developing models to support social learning in the domain of environmental management is a kind of participatory approach that can be very time-consuming (Barreteau et al. 2014). This is related to the wickedness of problems such as those encountered in environmental management, which are characterized by a high degree of scientific uncertainty and a profound lack of agreement on values (Balint et al. 2011). To confront such complex policy dilemmas requires developing learning networks (Stubbs & Lemon 2001) of stakeholders to create a cooperative decision-making environment in which trust, understanding, and mutual reliance develop over time.
- 8.7 Scientists are part of the group of people. To make the most of collaboration between them and other stakeholders, their role needs to be redefined towards scientists as stakeholders and participants, as has been suggested by Ozawa (1991), rather than scientists as an objective third party.
- 8.8 Using models to promote social learning implies a need to recontextualize uncertainty in a broader way — that is, relative to its role, meaning, and relationship with participants in decision-making (Brugnach et al. 2008). Instead of trying to eliminate or reduce uncertainties by collecting more data, the post-modern view of science (Funtowicz & Ravetz 1993) suggests that we should try to illuminate them (one of the purposes mentioned by Epstein in his 2008 paper), for their better recognition and acceptance by stakeholders who are then enabled to make informed collective decisions<sup>33</sup>.
- 8.9 Ambiguity is a distinct type of uncertainty that results from the simultaneous presence of multiple valid, and sometimes conflicting, ways of framing a problem. As such, it reflects discrepancies in meanings and interpretations. Under the presence of ambiguity, it is not clear what problem is to be solved, who should be involved

in the decision processes or what is an appropriate course of action (Brugnach & Ingram 2012). The co-design of a model may be a way to explore ambiguity among stakeholders through a constructivist approach (Salliou et al. 2017).

## Example

- 8.10** To mitigate the looming conflict in land use between government agencies seeking to rehabilitate the forest cover in upper watersheds of northern Thailand and smallholders belonging to ethnic minorities exploiting grasslands/fallows with an extensive cattle-rearing system a participatory modelling approach was applied. This involved an iterative series of collaborative agent-based simulation activities that was implemented in Nan Province (Dumrongrojwatthana et al. 2011). The exchange and integration of local empirical knowledge on the vegetation dynamics and extensive cattle rearing system was fostered by the co-design of the interactive simulation tools used to explore innovative land use options consistent with the dual interests of the parties in conflict. Herders requested the researchers to modify the model to explore new cattle management techniques, especially rotations of 'ruzi grass' (*Brachiaria ruziziensis*) pastures. These collaborative modelling activities, implemented with a range of local stakeholders (some of whom initially refused to even meet the herders), established a communication channel between herders, foresters and rangers. The dialogue led to an improved mutual understanding of their respective perceptions of land-use dynamics, objectives and practices. The improvement of trust between the villagers and the forest conservation agencies was also noticeable and translated into the design of a joint field experiment on artificial pastures.

## Risks

- 8.11** Finding the right way to engage stakeholders in co-designing a model of a complex system is tricky. Using easily accessible tools like role-playing games (Bousquet et al. 2002) can be better suited to grassroots people than for policy-makers (who tend to discard them as not serious enough). On the other hand, using conceptual designs and computer-implemented simulations with participants who did not receive any formal education may prove to be difficult. Beyond the risk of not selecting appropriate tools to support the co-design, there is also a risk that the models remain too generic and too abstract to get the interest of participants concerned by real-life problems. On the other hand, if one sticks to some particularities of a specific situation without stepping back sufficiently may not be enough to facilitate constructive discussions.
- 8.12** Without paying attention to power asymmetries, there is a risk that the most powerful stakeholders will have greater influence on the outcomes of the participatory modelling process than marginalized stakeholders.
- 8.13** In the same line, when there is an initial demand expressed by one of the stakeholders, there is a risk that the modellers bias the implementation of the process to fit their own perspectives (manipulation). As with any purpose, there is a danger that models are used to further the interests of a particular set of stakeholders over others, but this is a well known issue in the art and science of this kind of modelling.
- 8.14** Finally, there is a danger that goodness for this purpose implies a model is useful for other purposes, such as explanation or prediction. Truly participating in the development of a model can be a powerful experience and can give those involved a false impression of a model's empirical support for other purposes.

## Mitigating measures

- 8.15** Barnaud & Van Paassen (2013) advocate a critical companion posture, which strategically deals with power asymmetries to avoid increasing initial power imbalances. The goal of this is to shift more power from the modellers to the stakeholders. This posture suggests that designers of participatory modelling processes, intending to mitigate conflicts, should make explicit their assumptions and objectives regarding the social context so that local stakeholders can choose to accept them as legitimate or to reject them.

## Some Confusions of Purpose

- 9.1** It should be abundantly clear by now that establishing a simulation for one purpose does not justify it for another, and that any assumptions to the contrary risk confusion and unreliable science. However, the field has

many examples of such confusions and conflations, so we will discuss this a bit further. It is true that a simulation model justified for one purpose might be used as *part* of the development of a simulation model for another purpose — this can be how science progresses. However, just because a model for one purpose *suggests* a model for another, does not mean it is a good model for that new purpose. If it is being suggested that a model can be used for a new purpose, it should be separately justified for this new purpose. To drive home this point further, we look at some common confusions of purpose. Each time some code is mistakenly relied upon for a purpose other than has been established for it.

1. **Theoretical exposition → Explanation.** Once one has immersed oneself in a model, there is a danger that the world looks like this model to its author. This is a strong kind of Kuhn's 'Theoretical Spectacles'<sup>34</sup>, and results from the intimate relationship that simulation developers have with their model. Here the temptation is to jump from a theoretical exposition, which has no empirical basis, to an explanation of something in the world. An example of this is in the use of Lotka-Volterra models of species communities. Their realism is not only unclear, but actually not an issue at all, because they are just an illustration of an idea (what happens if, at equilibrium, abundance of species A affects growth rate of species B in that and that way?). This can be wrongly interpreted as an explanation of observed phenomena. The models may play a key conceptual role but miss out too much to be directly empirical. A simulation can provide a way of looking at some phenomena, but just because one can view some phenomena in a particular way does not make it a good explanation. Of course, one can form a hypothesis from anywhere, including from a theoretical exposition, but it remains only a hypothesis until it is established as a good explanation as discussed above (which would almost certainly involve changing the model).
2. **Description → Explanation.** In constructing a simulation for the purpose of describing a small set of observed cases, one has deliberately made many connections between aspects of the simulation and evidence of various kinds. Thus, one can be fairly certain that, at least, some of its aspects are realistic. Some of this fitting to evidence might be in the form of comparing the outcomes of the simulation to data, in which case it is tempting to suggest that the simulation supports an explanation of those outcomes. The trouble with this is twofold: (a) the work to test *which* aspects of that simulation are relevant to the aspects being explained has not been done; (b) the simulation has not been established against a range of cases — it is not general enough to make a good explanation.
3. **Explanation → Prediction.** A simulation that establishes an explanation traces a (complex) set of causal steps from the simulation set-up to outcomes that compare well with observed data. It is thus tempting to suggest that one can use this simulation to predict this observed data. However, the process of using a simulation to establish and understand an explanation inevitably involves iteration between the data being explained and the model specification — that is the model is fitted to that particular set of data. Model fitting is not a good way to construct a model useful for prediction, since it does not distinguish between what is essential for the prediction and the 'noise' (what cannot be predicted). Establishing that a simulation is good for prediction requires its testing against unknown data several times — this goes way beyond what is needed to establish a candidate explanation for some phenomena. This is especially true for social systems, where we often cannot predict events, but we can explain them after they have occurred.
4. **Illustration → Theoretical exposition.** A neat illustration of an idea suggests a mechanism. Thus, the temptation is to use a model designed as an illustration or playful exploration as being sufficient for the purpose of a theoretical exposition. A theoretical exposition involves the extensive testing of code to check the behaviour and the assumptions therein, an illustration, however suggestive, is not that rigorous. For example, it may be that an illustrated process is a very special case and only appears under very particular circumstances, or it may be that the outcomes were due to aspects of the simulation that were thought to be unimportant (such as the nature of a random number generator). The work to rule out these kinds of possibility is what differentiates using a simulation as an illustration from a theoretical exposition.
5. **Illustration → Prediction.** In 1972 a group of academics under the auspices of 'The Club of Rome' published a systems dynamics model that illustrated how laggy feedback between key variables could result in a catastrophe (in terms of a dramatic decrease in the variable for world population) (Meadows et al. 1972). By today's standards this model was very simple and omitted many aspects, including any price mechanism and innovation<sup>35</sup>. Unfortunately, this stark and much needed illustration was widely taken as predictive and the model criticised on these grounds<sup>36</sup>. The book does not stop asserting that the purpose was illustrative, "e.g. ... *in the most limited sense of the word. These graphs are not exact predictions of the values of the variables at any particular year in the future. They are indications of the system's behavioural*

*tendencies only*” (Meadows et al. 1972, pp. 92–93). However, the book also often conflates talk about the variables in their model and what those variables represent, e.g. “*Our world model was built specifically to investigate five major trends of global concern — accelerating industrialization, rapid population growth, widespread malnutrition, depletion of non-renewable resources, and a deteriorating environment*” (p. 21). It does claim that the model was not purely theoretical, since it says “We feel that the model described here is already sufficiently developed to be of some use to decision-makers...” (p. 22), however the way in which it would be of use to decision-makers was not made clear. This was an illustration of a possible future and not a prediction<sup>37</sup>.

6. **Social learning → Prediction.** Having been involved in a companion modelling process implemented in the National Park of Cevennes in the South of France (Etienne et al. 2003), a group of farmers, who had been sensitized to the process of pine encroachment of their pastures through participating in role-playing game sessions, decided to engage into an incentivized collective action plan. To specify the management plan, they requested to use as a decision support tool the agent-based model that had been designed to support the role-playing game sessions. The model was clearly not meant for such a use. A new version of the model was then designed, with a refined spatial resolution.

- 9.2 There is a natural progression in terms of purpose attempted as understanding develops: from illustration to description or theoretical exposition, from description to explanations and from explanations to prediction. However, each stage requires its *own* justification and probably a complete re-working of the simulation code for this new purpose. It is the lazy assumption that one purpose naturally or inevitably follows from another that is the danger.

## Modelling strategies

- 9.3 As well as a declared purpose for a model, most modellers will also employ a number of *modelling strategies* to guide the model development to achieve this purpose. A modelling strategy is an ‘rule of thumb’ that can be used as a guide as to *how* to develop the model. It is an informal meta-rule for the modelling processes, a story to help the modellers cut down the overwhelming number of choices presented to them to something more manageable. A modelling strategy is more about the process of model *development*, the modelling purpose is more about its *justification*. This is essentially the same as the classic distinction in the philosophy of science between the ‘context of discovery’ and the ‘context of justification’ (Reichenbach 1938). Examples of modelling strategies include “KISS” (Keep It Simple Stupid) and “KIDS” (Keep It Descriptive Stupid) (Edmonds & Moss 2005).
- 9.4 Although there has been much debate about these strategies (e.g. Medawar 1967; Loehle 1990; Edmonds & Moss 2005; Grimm et al. 2005; Edmonds 2010; Sun et al. 2016; Jager 2017), in this paper we want to concentrate on model purpose. Since, usually, we are not privy to the process of model development but only the justification of models we often do not know which strategies people are actually using<sup>38</sup>, but we feel we should be explicitly and clearly told the modelling purpose. Also modelling strategy is more of a matter of personal style. However, distinguishing different purposes illuminates the debate on strategy so we discuss how different strategies might be appropriate for different purposes in the Appendix.

## Using Models with Different Purposes Together

- 10.1 Clearly one sometimes needs to be able to achieve more than one of the above purposes with respect to some phenomena. In such circumstances one might be tempted to imagine a model that mirrors reality so well that it can be used for many purposes (e.g. to both predict and explain). In the above, we have argued that, even if that were achievable, you would need to justify the model for each purpose separately.
- 10.2 However, we *further* suggest that this is usually better done with separate, but related models. The reason for this should have become clear during the discussion as to what modelling strategies might be helpful to achieve each purpose — each purpose has different concerns and hence different strategies are helpful. Even if one had a model of something as simple as an ideal gas, with a different entity for each gas particle, it might not be a very practical model for prediction<sup>39</sup>. Rather one would probably use the gas laws that connect volume, temperature and pressure that approximate this individual-based model.
- 10.3 Cartwright (1983) argues that, in physics, there is a distinction between explanatory and phenomenological laws. That is, in the above language, the models that explain and the models that predict are different. However,



in a mature science, the relationship between these different models is well understood. Cartwright describes some examples where an explanatory theory is used as the starting point for generating a number of alternative predictive models and the predictions are only trusted when the alternate routes to that prediction roughly agree.

- 10.4 In the social sciences, the phenomena are more complex than those usually tackled in physics. For this reason, we should expect there to be many different models that achieve different purposes in respect to some phenomena, different versions of these models for different assumptions one might make, and maybe different versions to address different aspects of those phenomena. Thus, we might expect that a whole cluster of different models might be used to get a better picture of any particular phenomenon, each with different purposes. The idea that we might have only one model that is adequate for one target is probably infeasible.
- 10.5 An example of using different models for different purposes is when models of different levels of abstraction are used to 'stage' abstraction. Following a KIDS kind-of strategy, one can start with a relatively complex or descriptive model and then seek to approximate this with simpler models that one can get a better theoretical understanding of. This was tried in Fieldhouse et al. (2016); Lafuerza et al. (2016b,a).
- 10.6 By keeping models for different purposes separately but defining/describing the relationship between them, we can get many of the advantages of combining models but with less chance of confusion of purpose.

## Conclusion

- 11.1 In Table 1 we summarise the most important points of the above distinctions. This does not include all the risks of each kind of model, but simply picks what we see as the most pertinent ones. You will notice that the features and risks are sometimes associated with validation and sometimes verification activities; this is because the relative importance of validation and verification (and indeed other activities such as interpretation) change dependent on the model purpose. If one is aiming at using a model for prediction, then validation is crucial (an essential feature), but if it is for theoretical exploration then verification is key.

Modelling Purpose	Essential features	Particular risks (apart from that of lacking the essential features)
<i>Prediction</i>	Anticipates <i>unknown</i> data	Conditions of application unclear
<i>Explanation</i>	Uses <i>plausible</i> mechanisms to match outcome data in a <i>well-defined</i> manner	Model is brittle, so minor changes in the set-up result in bad fit to explained data; bugs in the code
<i>Description</i>	Relates directly to evidence for a set of cases	Unclear provenance; over generalisation from cases described
<i>Theoretical exposition</i>	Systematically maps out or establishes the consequences of some mechanisms	Bugs in the code; inadequate coverage of possibilities
<i>Illustration</i>	Shows an idea clearly as a particular example	Over interpretation to make theoretical or empirical claims; vagueness
<i>Analogy</i>	Provides a way of thinking about something; gives insights	Taking it seriously for any other purpose
<i>Social learning</i>	Facilitates communication or agreement	Lack of engagement; confusion with objective modelling

Table 1: A brief summary of the discussed modelling purposes

- 11.2 As should be clear from the above discussion, being clear about one's purpose in modelling is central to how one goes about developing, checking and presenting the results. Different modelling purposes imply different risks, and hence different activities to avoid these. If one is intending the simulation to have a public function (in terms of application or publication), then one should have taken action to minimise the particular risks to achieving its purpose, so that others can rely on the model for this purpose (at least to some extent).
- 11.3 This does not include private modelling, whose purpose is maybe playful or exploratory and done only between consenting adults (or alone). However, in this case one should not present the results or model as if they have achieved anything more than an illustration, otherwise one risks wasting other people's time and attention. If

one finds something of value in the exploration it should then be re-worked so it can be justified for a particular purpose to be sure it is worth public attention.

- 11.4** This is not to say that models that have been published without a clear purpose should not have been published, but that now the field has matured we should aim for higher and more scientifically productive standards – one of these standards is that an audience should know under what purpose a model is being justified, so they know how to judge it (and hence in what ways they might rely on it in their work). Perhaps journals should encourage authors to clarify the purpose of their models from the beginning of the submission process, so that editors, reviewers and (later) readers will know how to critique the research and avoid some of the potential misunderstandings.

*A confused, conflated or unclear modelling purpose leads to a model that is hard to check, can create misleading results, and is hard for readers to judge — in short, it is a recipe for **bad science**.*

## Acknowledgements

This manuscript was written by Bruce Edmonds and Christophe le Page, with suggestions and additions by the other authors. The authors wish to thank all those with whom they have discussed these issues, who are far too numerous for us to remember. However, these do include Scott Moss, Alan McKane, Juliette Rouchier, David Hales, Guillaume Deffuant, Mario Paolucci, Rosaria Conte, Eckhart Arnold and Dunja Šešelja. Also all those at the workshop on Validation for Social Simulation held in Manchester in the summer of 2014 (including: Nigel Gilbert, Mike Batty, Jeff Johnson, and Rob Axtell) whose contributions and discussions helped formulate these ideas. If you are not on this list, you should have argued harder. This manuscript has been handled by the former journal editor, Nigel Gilbert.

## Appendix: About Modelling Strategies

A modelling strategy is an informal ‘rule of thumb’ that might be used to guide model development and checking. Thus, the following discussion is more speculative than the rest of the paper, which is why it is relegated to an Appendix. Such strategies are not something that can be proved right or wrong, merely more or less helpful in the experience of individual modellers. However, as already noted, discussion about these strategies have pervaded the literature, so it is worth saying what light distinguishing between different model purposes has on this debate.

Some of these strategies are as follows.

- **Keep it Simple Stupid (KISS)**. This is the strategy of starting as simple as one can get away with and then only adding complication when necessary. Thus, one starts with relatively simple models and develops from there. This is a well-established principle in engineering since it maintains the maximum control and understanding of the system as it develops, and hence reduces the possibility of making mistakes. A particular example of KISS is the **TAPAS** strategy, namely “Take a Previous Model and Add Something” (Frenken 2006) where you start from a previously published model<sup>40</sup>.
- **Minimise the Number of Parameters (MNP)**. This is where one has as few parameters for the model as possible. This is slightly different from KISS in terms of both purpose and practice. Under MNP, you can have complex processes as long as the number of parameters is small because the point is to reduce the space of output behaviours. A random number generator may be implemented using a complex process, but has only one parameter (its seed). A linear equation with many variables is simple but has many parameters.
- **Minimise the Number of Free Parameters (MNFP)**. A free parameter is one that is not determinable using measurement from the target of modelling, even in principle. These might be scaling factors (e.g. relating different time scales) or represent unknown aspects (e.g. the parameters of some learning algorithm that an agent uses). Having fewer free parameters makes observed data difficult to ‘fit’ just by adjusting parameters. Free parameters are sometimes estimated by fitting a model to in-sample data, but then one is not clear which aspects of that data are being used and that in-sample data becomes (in effect) a huge set of parameters.

- **Keep it Descriptive Stupid (KIDS).** This is the strategy of starting from the evidence<sup>41</sup> — if there is empirical evidence for some aspect of the phenomena of interest then you include this in your model by default. It is important that this is not just how one *thinks* things are — otherwise this strategy can degenerate into MTMMLTP (below). One then starts the model exploration from this point, and then maybe tries the model without some of the elements and seeing if it makes any significant difference to the model behaviour. However, as with the KISS strategy, one might equally find one needs even more detail. This is described in (Edmonds & Moss 2005).
- **Make the Model More Like the Phenomena (MTMMLTP<sup>42</sup>).** This is the strategy of making the model more like how we *think* things are. In other words, to use the correspondence view of modelling (discussed at the beginning) as a guide to model development. We suspect that most of us use this strategy to some extent, whether we are aware of it or not. This, we claim, is the source of much misunderstanding — just because we *think* of a model as being like some part of reality does necessarily not make the model good for any particular purpose. The strength of this strategy depends critically on how much our view of how things are is informed by empirical evidence — this is the important difference between this and the KIDS strategy.
- **Enhancing the Realism of Simulation (EROS).** This might seem to be a specific sub-case of MTMMLTP, since it advocates expanding the cognitive models used in simulations to be more psychologically plausible (Jager 2017). However, both Wander Jager and Rosaria Conte operationalize this by making the cognitive models in their simulations consistent with some aspect of psychological theories — thus this is more of theoretical constraint of one subject upon another rather than psychological realism. This is obviously useful if you want an explanation to be in terms of such psychological theories, but are ultimately only as strong as these theories.
- **Keep It a Learning Tool (KILT).** This is the strategy of maximising the relevance and accessibility of a model so that participants can engage with the model and be willing to help specify or improve it (Le Page & Perrotton 2017). This is somewhat different to the other strategies here, since it is tied in with the purpose of social learning. Social learning is a case where the process of model building is its point, and not its subsequent justification.

These strategies are not something that can ever be proved or disproved, but rather only shown as helpful or otherwise during the processes of modelling — especially the extent to which a model's purpose is achieved and demonstrated using them. To some extent, these are a matter of taste and personal style, but we argue that they are also differently relevant to different modelling purposes. We consider each purpose in turn.

## Prediction

For the bare purpose of prediction it actually does not matter why your model predicts, just that it does<sup>43</sup>. However, one does need some understanding of the code, because in practice the only way to find a predictive model is a process of trial and error, adapting the code until it predicts — if one does not understand the code it is hard to know how to adapt this to improve it for the next trial. Thus, KISS plays a role here. What is more important to this purpose is (a) being able to set the parameters correctly for a case where one wants to predict and (b) that the sensitivity of the key outcomes to those parameters can be established so that one knows to the likely accuracy of the prediction. For (a), MNFP is important so the model can be set up correctly for a prediction. For (b), MNP is important so that the sensitivity (including joint sensitivity) of the predictions to the parameters can be established using a Monte Carlo approach (or similar). If there is very reliable evidence for key parts of the process one is simulating (e.g. the electoral college method of electing a president of the US) then it might be sensible to fix that in your code. Thus, KIDS does have a small role here. MTMMLTP is not necessarily relevant here and could result from adding more parameters, which would be counter-productive.

## Explanation

When explaining some measured phenomena using a model, one does so in terms of the processes and structures that are built into the model. If those process and structures are not strongly related to those observed one gets an explanation in terms that are not so relevant to what is being explained. Rather for explanation, it is vital that those aspects of the phenomena that we want the explanation to be in terms of, are in the model.

Explanation differs from that of analogy because it has a well-defined relationship between the model and evidence of what is observed, so that the KIDS strategy is very important here. It is desirable to have few free parameters, as otherwise there is a danger that the ‘fit’ to observed outcomes was due to these having special values, so MNFP plays a role. Social phenomena are often quite complex so explanatory models are often complex (hence KISS is not relevant) and often have many parameters, but these should be based on evidence so there might be many of these, so that MNP is not so relevant here. MTMMLTP might lead one to include aspects for which there is no direct empirical evidence, which would increase the number of free parameters, so MTMMLTP could also be counter-productive here. However, if one wants some empirical outcomes explained in terms of psychological theories then EROS is appropriate.

## Description

In a description you are just trying to record the observed aspects of some phenomena for latter use and inspection. Ideally you should not be projecting what you ‘think’ is there but focussing on what is observed — i.e. what there is evidence for. Thus it is KIDS and not MTMMLTP that is more relevant here. The number of parameters (MNP) is not important, but having lots of free parameters is an indication you are including unobserved aspects into your description, so MNFP is. A description should be as simple or as complex as what is being described so KISS is not relevant here either.

## Theoretical exposition

When trying to systematically characterise the behaviour of a model, there are essentially two techniques:

1. to systematically map out the space of behaviours by running the model many times
2. to approximate the behaviour with a simpler model (e.g. an analytic model, a social network analysis or a simpler simulation).

The *first* of these becomes more difficult with an increasing number of parameters that need to be mapped. Thus, MNP is core here. Whether the parameters are free or not is not really relevant to a theoretical exposition (so MNFP is not very important here). The second of these is harder the more complicated and complex the original model is, thus KISS is also key. Although theoreticians may have an eye to the eventual applicability of their work, the relationship of their model to what is observed is not important. Whilst evidence from observed systems might be a helpful guide to whether one is on the right track, it is well known that trying to add too many features just because that is how one thinks of a system can be fatal to a good exposition (thus MTMMLTP is unhelpful here).

## Illustration

A good illustration is as simple as possible to capture what is needed. Thus, KISS has to be the over-riding principle here and it probably does not have many parameters. It can help if the illustration is plausible (MTMMLTP). An illustration is often of an idea rather than an observed process, so KIDS does not usually come into this.

## Analogy

A good analogy is relatively simple (KISS) but is flexible as to its relationship with what is observed. Indeed some of the point of an analogy is that it can be applied to completely unforeseen situations (so not KIDS). It helps if it is plausible (MTMMLTP). Since analogies are often used in social learning KILT is relevant here (although we mediate so naturally using analogies that there is less need for a ‘science’ of this).

## Social learning

Regarding the suitability of the tools and the degree of model complication (Sun et al. 2016), Le Page & Perrotton (2017) advocate using an intermediate level of stylized but empirically based models and proposed the motto Keep It a Learning Tool (KILT) to stress the importance of sticking to the announced purpose of promoting social

learning among participants. The KISS approach can initiate the process with an over-simplified stylized yet empirically grounded model that enables tackling the complexity of the target socio-ecosystem with a tool that has the status of a sketch. It provides the main features of the final version; however, it is clearly unfinished: there remains an important work of progressive shaping and improvement so that it acquires its final form and becomes usable with people who were not involved in its design. The crudeness of the initial version of the model is actually a strategy to stimulate the involvement of the participants targeted as potential co-designers.

This strategy includes elements of MTMMLTP since aspects of what is being represented in the model (those assumed by all those involved) need to be recognisable by the stakeholders, especially if they structure the more contentious elements (e.g. geographical layout, or seasons of the year) — there does not *have* to be evidence for these since it is the perception of participants that counts, but there might be. The processes involved do need to be comprehensible enough to be understood and criticisable by the participants, so KISS does play a role here but not a central one. On the other hand, having many parameters (free or otherwise) can be an advantage since these might be adjusted to enhance the plausibility or ‘realism’ of a simulation, so MNP is not helpful here.

## Summary of strategies

Table 2 summarises the above discussion in terms of the rough importance of the various strategies for achieving the various model purposes. This is **not** definitive, since strategy is partly a matter of taste. It can not be a prescription but a rule of thumb that might be helpful.

Modelling Purpose	Probably Important	Can be Helpful	Possibly Unhelpful
<i>Prediction</i>	MNFP, MNP	KISS, KIDS	MTMMLTP
<i>Explanation</i>	KIDS	MNFP, EROS	MTMMLTP
<i>Description</i>	KIDS, EROS	MTMMLTP, MNFP	KISS
<i>Theoretical exposition</i>	MNP, KISS	KIDS	MTMMLTP
<i>Illustration</i>	KISS	MTMMLTP, MNP	EROS
<i>Analogy</i>	KISS	MTMMLTP, KILT	KIDS, EROS
<i>Social learning</i>	KILT	MTMMLTP, KISS	MNP

Table 2: Discussed relevance of strategies for the development of models with respect to the various purposes  
heck your code and try to refute the explanation using carefully designed experiments with the model.

## Notes

<sup>1</sup>Zeigler et al. (1976) describes this as not one-one but as a homomorphism (a one-many) relationship so that parts of the model correspond to sets of elements in what is observed, but this is effectively the same.

<sup>2</sup>With the exception of the purpose of description where a model is *intended* to reflect what is observed.

<sup>3</sup>We prefer the idea of a tool to that of a mediator (Morgan et al. 1999) since a model as a mediator is too vague and frustrates any assessment of goodness for purpose (except for the purpose of social interaction where people are good at assessing the quality of communication between themselves, and hence how useful a model was in facilitating this).

<sup>4</sup>Thiele & Grimm (2015) recommend to re-implement existing models, or parts of them, instead of just using the code.

<sup>5</sup>See discussion about this problem in Bell et al. (2015).

<sup>6</sup>We are not ruling out the possibility of re-usable model components in the future using some clever protocol, we have now read many proposals improving the re-usability of model code, but not seen any cases where this works for anyone except for the team making these proposals.

<sup>7</sup>“Thinking about things analogically” would be a more precise title but that is too long.

<sup>8</sup>Short alternatives for this might be social learning, collaborative learning or the French word *concertation*. We went with the broadest term.

<sup>9</sup>In ecology, for example, they are the most common model purposes, together with “theoretical exploration”.

<sup>10</sup>Validation is when one checks a model against its real world data, verification is when one checks that the simulation does correspond to what we intended (Ormerod & Roswell 2009).

<sup>11</sup>It would not really matter even if the code had a bug in it, if the code reliably predicts (though it might impact upon the knowledge of *when* we can rely upon it or not).

<sup>12</sup>Where model B may be a random or null model but also might be a rival model.

<sup>13</sup>This kind of prediction is somewhat problematic, since it is often only apparent after additional empirical work has revealed it to be correct and thus might be better characterized as a kind of explanation. However, it is the most commonly confirmed kind of prediction for complex simulation models, and it is possible that some predictions were put in print before they were checked for, so we included it.

<sup>14</sup>In ecological modelling, for example, it is common practice to plot model output versus data and refer to this as “predicted vs. observed”, although in virtually all cases those “predictions” are based on heavy calibration, i.e. tweaking.

<sup>15</sup>There is a sense in which any model output can be re-cast as a conditional prediction (Boschetti et al. 2011) but in many cases (e.g. where a model is not empirically based) the conditions are so unlikely as to be meaningless. Thus we are not convinced that this interpretation of models is very useful — many models are simply not intended to predict.

<sup>16</sup>To be precise, some people have claimed to predict various social phenomena, but there are very few cases where the predictions are made public before the data is known and where the number of failed predictions can be checked. However Polhill (2018) argues for the importance of prediction, despite the difficulties. See Stillman et al. (2014) for a discussion of this in ecology.

<sup>17</sup>Good practice might be to separate the training data (that informs your modelling), calibration data (where free parameters are estimated), model selection data (for whether one would pick that model to publish it), and performance assessment data (which gives an idea of its objective performance).

<sup>18</sup>On top of these are those common to all kinds of complex system, including chaos, second-order emergence etc.

<sup>19</sup>This is, in general difficult to tell without also a good explanation of why a model predicts well. However, as we pointed out the lack of a good explanation does not necessarily stop prediction being possible, as with some “deep learning” models trained on big data.

<sup>20</sup>In a mature science one will have both explanatory and predictive models, as well as how these are related, but as science develops one might come before the other, which is one of the reasons it is important to distinguish these.

<sup>21</sup>We are being a little disparaging here, it may be that these have a definite meaning in terms of relating different scales or some such, but too often they do not have any clear meaning but just help the model fit stuff.

<sup>22</sup>Whilst ideally, processes that are not thought essential to the explanation should not be included in models designed for the purpose of supporting explanations, in practice it is often impossible not to include some, if only to make the simulation run.

<sup>23</sup>The core of this was essentially François’ doctoral research.

<sup>24</sup>To be precise, it does assume there are discrete entities or objects and that there are processes within these that can be represented in terms of computations, but these are not very restrictive assumptions.

<sup>25</sup>Thus making modelling path- and culture-dependent.

<sup>26</sup>As in so called ‘Model to Model’ analysis (Hales et al. 2003).

<sup>27</sup>In the sense of not being vulnerable to being shown to be wrong later.

<sup>28</sup>Later, it stimulated a growing category of illustrative counter-example models of unintended macro consequences (Hegselmann 2017).

<sup>29</sup>Lotka described modelling as translating a verbal idea into an “animated object”, which has a life of its own and can surprise.

<sup>30</sup>To be exact, the relationship is between some part of the understanding of the working of a simulation to the domain of application.

<sup>31</sup>Indeed the work spawned a whole industry of papers doing just such an exploration.



<sup>32</sup>For example, the spontaneous outbreaks of non-hostility across the trenches in World War I.

<sup>33</sup>Which might include a recognition of the need to collect more data.

<sup>34</sup>Kuhn (1962) pointed out the tendency of scientists to only see the evidence that is coherent with an existing theory — it is as if they have ‘theoretical spectacles’ that filter out other kinds of evidence.

<sup>35</sup>The model was heavily criticised by economists from soon after its publication, on the grounds that it did not allow feedback from the price mechanism which would in principle encourage the conservation of scarce resources and provide an incentive to innovate.

<sup>36</sup>For example <http://foreignpolicy.com/2009/11/09/the-dustbin-of-history-limits-to-growth/>, which mocks the quality of predictions.

<sup>37</sup>We are aware that there are some that argue that, in fact, it was pretty good as a predictive model (e.g. <https://www.theguardian.com/commentisfree/2014/sep/02/limits-to-growth-was-right-new-research-shows-were-nearing-collapse>), but the model only ‘predicts’ in some and not other aspects over a short time frame and in the absence of major turning points in the data.

<sup>38</sup>Grimm et al. (2014) argue that we should record and make public the process of model development in ‘TRACE’ documents.

<sup>39</sup>Of course, it is possible that an individual-based model might be useful for predicting other things, such as how much gas might leak through a particular shaped hole.

<sup>40</sup>This has the added advantage of making it easier to publish in some fields, but that is beside the point here.

<sup>41</sup>We must admit that the acronym is a bit of a misnomer since here it is evidence rather than representation that is key to it, however KIEB (Keep It Evidence Based) was not such a good acronym and would have ruined the title of Edmonds & Moss (2005).

<sup>42</sup>OK, we admit it — we made this acronym long and unpronounceable so that it would not catch on.

<sup>43</sup>It is important to know when and where a model works as a predictive model, e.g. which of two predictive models one should use for any particular situation. This is why some knowledge about the conditions of application of predictive models is important — even if this is just a rule of thumb. Sometimes such knowledge can be provided by a good explanation, but sometimes this is the result of past trial and error.

## References